Stochastik

Zufallsgrößen, Baumdiagramme und Vierfeldertafeln

Handreichung zur Notation in schriftlichen Arbeiten
Zusammenstellung des Gymnasium Sulingen

Fassung 06/2011

Vorwort

Diese Handreichungen werden für den Mathematikunterricht am Gymnasium Sulingen als verbindliche Minimaldokumentationen zur Lösung von Aufgaben aus dem Bereich der Stochastik angesehen. Für die Dokumentation der Lösungswege im schriftlichen Abitur können daher diese Informationen hilfreich sein, um einen strukturierten Lösungsweg darzustellen.
Aufgaben in schriftlichen Prüfungen sind operationalisiert, d. h. in der Aufgabenstellung werden Verben benutzt, die auf einen genau vorgegebenen Arbeitsauftrag schließen lassen. An dieser Stelle sollen die beiden Operatoren *berechnen* und *bestimmen* klar differenziert werden. In den Vorgaben des Niedersächsischen Kultusministeriums aus dem Jahr 2009 werden die beiden Operatoren folgendermaßen definiert:

Berechnen

Es wird von der Schülerin bzw. dem Schüler erwartet, dass **die Ergebnisse von einem Ansatz ausgehend gewonnen werden**. Beispiele hierfür sind unter anderem: „Berechnen Sie die Wahrscheinlichkeit des Ereignisses.“, „Berechnen Sie den Flächeninhalt…“ oder „Berechen Sie die größtmögliche Höhe…“. Es sind hierfür alle Werkzeugebenen zulässig. Durch Zusätze zum Operator *berechnen* sind Einschränkungen oder weitere Vorgaben möglich (z. B. berechnen Sie algebraisch, berechnen Sie numerisch usw.). Es kann bei der Verfügbarkeit von GTR bzw. CAS im Einzelfall die Darstellung eines Lösungsweges oder einer Lösung so gefordert werden, das diese auch ohne den Einsatz der eingesetzten Technologien nachvollziehbar ist.

Bestimmen / Ermitteln

Es wird von der Schülerin bzw. dem Schüler erwartet, dass **ein möglicher Lösungsweg dargestellt wird und das Ergebnis formuliert wird**. Beispiele hierfür sind unter anderem: „Ermitteln Sie den Schnittpunkt.“ oder „Bestimmen Sie aus diesen Werten die Koordinaten der beiden Punkte.“. Es sind hierfür alle Werkzeugebenen zulässig. Durch Zusätze zum Operator *berechnen* sind Einschränkungen oder weitere Vorgaben möglich (z. B. berechnen Sie algebraisch, berechnen Sie numerisch usw.). Es kann bei der Verfügbarkeit von GTR bzw. CAS im Einzelfall die Darstellung eines Lösungsweges oder einer Lösung so gefordert werden, das diese auch ohne den Einsatz der eingesetzten Technologien nachvollziehbar ist.

Entscheidend sind dabei die **hervorgehobenen** Details. Die Fachgruppe Mathematik am Gymnasium Sulingen erwartet daher von den Schülerinnen und Schülern bei der Angabe des Operators *berechnen* einen mathematisch vollständigen Lösungsansatz, der, sofern dies nicht anders durch Einschränkungen angegeben wird, mit allen Hilfsmitteln gelöst werden kann. Ist durch den Operator *bestimmen* der Arbeitsauftrag definiert worden, so ist es erlaubt einen beliebigen Lösungsweg zu dokumentieren, der nachvollziehbar sein muss, aber nicht zwingend einen mathematischen Ansatz beinhaltet. Auch hierfür ist, wenn nicht anders in der Aufgabenstellung angegeben, jedes Hilfsmittel zulässig. Entscheidend ist im Falle von *bestimmen* eine inhaltlich korrekte Angabe des erhaltenen Ergebnisses.

Grundsätzliches Vorgehen

Zu jedem dokumentierten Lösungsweg gehören vier maßgebliche Abschnitte, die in der Dokumentation wiedergefunden werden sollen:

(1) Informationen über gegebenene Größen,
(2) Informationen über gesuchte Größe im Aufgabenkontext,
(3) Vorgehensweise bzw. mathematischer Ansatz,
(4) Interpretation und Argumentation anhand erhaltener Ergebnisse.

Im nachfolgenden Teil werden Musteraufgaben mit Lösungen dargestellt. Sämtliche Bestandteile einer Musterlösung werden für die vollständige Dokumentation bei Aufgabenstellungen mit dem Operator *berechnen* von den Schülerinnen und Schülern erwartet.

Kommentare, die nicht explizit zur Dokumentation des Lösungsweges gehören, werden grau hinterlegt.

Bestimmung von Wahrscheinlichkeiten bei binomialverteilter Zufallsgröße

Aufgabenstellung

In verschiedenen Ländern wurden in den letzten Jahren immer häufiger Wohngebiete gebaut, in denen (fast) ausschließlich Menschen über 55 Jahren leben; auch in Uranien werden solche „Seniorenstädte“ geplant.
Aufgrund von Erhebungen weiß man, dass 40% der Uranier über 55 Jahre (im Folgenden kurz als Senioren bezeichnet) Interesse für eine solche Seniorenstadt bekunden.
Bestimmen bzw. berechnen Sie die Wahrscheinlichkeit dafür, dass sich unter 20 zufällig ausgewählten Senioren

(5) genau 8 Interessenten befinden,
(6) weniger als 4 Interessenten befinden.

Hinweis: Es sollen Aufgaben bearbeitet werden, die das Verständnis der binomialverteilten Zufallsgrößen (in Kursen mit erhöhtem Niveau auch normalverteilte Zufallsgrößen) fördern, d. h. es sollen Aufgaben gestellt werden, die eine Begründung einfordern, warum es sich um genau diesen Typ einer Wahrscheinlichkeitsverteilung handelt.

Erwartungshorizont

X: Anzahl der interessierten Senioren
X ist binomialverteilt.

gegeben: $n = 20$, $p = 0{,}4$, $k = 8$ bzw. $k < 4$.

gesucht: (1) $P(X = 8)$ und (2) $P(X < 4)$

zu (1): $P(X=8)=\binom{20}{8}\cdot 0{,}4^{8}\cdot 0{,}6^{12}\approx 17{,}97\%$

binomPdf(20, 0.4, 8) 0.1797

$P(X=8)\approx 17{,}97\%$

Antwort: Die Wahrscheinlichkeit unter 20 zufällig ausgewählten Senioren genau 8 Interessenten zu finden, beträgt etwa 17,97%.

Hinweis: Die Verwendung des Taschenrechners zur rechnerischen Ermittlung des Zahlenwertes der Wahrscheinlichkeit wird in zwei Teilschritten dokumentiert. Auf der linken Seite findet sich die **Taschenrechnereingabe** (mit nach Möglichkeit weitestgehend mathematischer Schreibweise von Termen) und auf der rechten Seite die **Taschenrechnerausgabe** nach der Berechnung durch das Gerät.
Eine sinnvolle Rundung der Taschenrechnerergebnisse ist im Regelfall Bestandteil der Aufgabenstellung und kann je nach Aufgabenkontext in der Dezimalstellenanzahl variieren.

zu (2): $P(X<4)=P(X=0)+\ldots+P(X=3)\approx 1{,}6\%$

alternativ: $P(X<4)=\sum_{k=0}^{3}\binom{20}{k}\cdot 0{,}4^{k}\cdot 0{,}6^{20-k}\approx 1{,}6\%$

binomCdf(20, 0.4, 0, 3) 0.016

$P(X<4)=P(X\leqslant 3)\approx 1{,}6\%$

Antwort: Die Wahrscheinlichkeit, dass unter 20 zufällig ausgewählten Senioren weniger als 4 Interessenten zu finden, beträgt etwa 1,6%.

Hinweis: Es ist auch denkbar, dass anstelle der exakten mathematischen Notation mit Binomialkoeffizienten und Summennotation die Taschenrechnerbefehle binomPdf und binomCdf als mathematische Funktion mit einer Unbekannten und zwei bzw. drei Parametern definiert werden.

$\text{binomPdf}(n,\ p,\ k):=\binom{n}{k}\cdot p^{k}\cdot\left(1-p\right)^{n-k}$ um $P(X=k)$ zu berechnen,

$\text{binomCdf}(n,\ p,\ a,\ b):=\sum_{k=a}^{b}\binom{n}{k}\cdot p^{k}\cdot\left(1-p\right)^{n-k}$ um $P(a\leqslant X\leqslant b)$ zu berechnen.

Bestimmung von bestimmten Wahrscheinlichkeiten über ein vollständiges Baumdiagramm

Aufgabenstellung

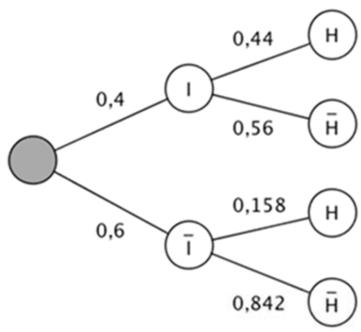
Eine Befragung ergab, dass 44% der Senioren, die Interesse haben, in der Seniorenstadt zu leben, ein Haustier besitzen. Andererseits haben 65% der Senioren, die ein Haustier besitzen, Interesse für die Seniorenstadt bekundet. Bestimmen Sie den prozentualen Anteil der Senioren, die

(1) an der Seniorenstadt interessiert sind und ein Haustier besitzen,
(2) ein Haustier besitzen.

Hinweis: Hier handelt es sich um bedingte Wahrscheinlichkeiten. Diese Aufgabe ist mithilfe eines Baumdiagrammes oder einer Vierfeldertafel lösbar. Ein vollständiges Baudiagramm enthält alle Pfadwahrscheinlichkeiten (siehe Beispiel unten). Die im Baumdiagramm verwendeten Abkürzungen für eintretende Ereignisse müssen definiert worden sein.

Erwartungshorizont

gegeben: H: besitzt Haustier, $\overline{\text{H}}$: besitzt kein Haustier,
I: Senioren haben Interesse, $\overline{\text{I}}$: Senioren haben kein Interesse.



gesucht: (1) $P(I \cap H)$ und (2) $P(H)$

Hinweis: Je nach Formulierung im Lösungsweg sind anstelle der mengentheoretischen Schreibweise als Schnittmenge auch Dokumentationen der Art P(I und H) oder P(„Senioren haben Interesse und besitzen ein Haustier“) denkbar und als gleichwertig zu betrachten.

zu (1): $P(I \cap H) = 0{,}4 \cdot 0{,}44 = 17{,}6\%$

Antwort: Der prozentuale Anteil der Senioren, die an der Seniorenstadt interessiert sind und ein Haustier besitzen, beträgt 17,6%.

zu (2): $$P(H) = \frac{P(I \cap H)}{P_H(I)} = \frac{0{,}176}{0{,}649} \approx 27{,}1\%$$

alternativ dürfen verwendet werden:
P(I∩H) = P(„Senioren haben Interesse und besitzen ein Haustier“),
$P_H(I)$ = P(„Senioren haben Interesse unter der Bedingung, dass sie ein Haustier besitzen“).

Antwort: Der prozentuale Anteil der Senioren, die ein Haustier besitzen, beträgt etwa 27,1%.

Hinweis: Falls ein Baumdiagramm gezeichnet bzw. eine Vierfeldertafel erstellt wurde, können zu berechnende Wahrscheinlichkeiten direkt abgelesen werden. Die Vierfeldertafel sollte den unteren Aufbau haben.

	Haustier	**kein Haustier**	**Summe**
Senioren haben Interesse	17,6%	22,4%	40%
Senioren haben kein Interesse	9,5%	50,5%	60%
Summe	27,1%	73,9%	100%

Bestimmung von bestimmten Wahrscheinlichkeiten über ein vollständiges Baumdiagramm

Aufgabenstellung

Laut Statistiken aus anderen Ländern sei das Alter der Menschen, die in diesen Seniorenstädten leben, normalverteilt mit $\mu = 65$ Jahren und $\sigma = 5$ Jahren. Ein Patient besucht eine Arztpraxis in der Seniorenstadt.
Berechnen Sie die Wahrscheinlichkeit,

(1) dass das Alter des Patienten um maximal 5 Jahre vom Erwartungswert abweicht,
(2) dass der Patient älter als 70 Jahre ist,
(3) dass der Patient jünger als 55 Jahre ist.
(4) Geben Sie ein Intervall (Altersspanne) an, in dem etwa 89% der Patienten liegen.

Erwartungshorizont

X: Anzahl der Senioren, die eine Arztpraxis besuchen
X ist binomialverteilt.

gegeben: $\mu = 65$; $\sigma = 5$; je nach Aufgabenstellung k
(1) $60 \leqslant k \leqslant 70$; (2) $k > 70$; (3) $k < 55$
(4) Die Wahrscheinlichkeit p des gesuchten Intervalls beträgt 89%.

gesucht: (1) $P(60 \leqslant X \leqslant 70)$
(2) $P(X > 70)$
(3) $P(X < 55)$

(4) Intervall für X

zu (1): $P(60 \leqslant X \leqslant 70) = \dfrac{1}{5\sqrt{2\pi}} \displaystyle\int_{60}^{70} e^{-\left(\frac{x-65}{5}\right)^2} dx \approx 68{,}3\%$

normCdf(60, 70, 65, 5) 0.683

$P(60 \leqslant X \leqslant 70) \approx 68{,}3\%$

Antwort: Die Wahrscheinlichkeit, dass ein Patient im Alter um 5 Jahre von μ abweicht, beträgt etwa 63,8%.

zu (2): $P(X > 70) = \dfrac{1}{5\sqrt{2\pi}} \displaystyle\int_{70}^{\infty} e^{-\left(\frac{x-65}{5}\right)^2} dx \approx 15{,}87\%$

normCdf(70, ∞, 65, 5) 0.1587

$P(X > 70) \approx 15{,}87\%$

Antwort: Die Wahrscheinlichkeit, dass ein Patient älter als 70 Jahre alt ist, beträgt etwa 15,87%.

zu (3): $P(X < 55) = \dfrac{1}{5\sqrt{2\pi}} \displaystyle\int_{-\infty}^{55} e^{-\left(\frac{x-65}{5}\right)^2} dx \approx 2{,}28\%$

normCdf($-\infty$, 55, 65, 5) 0.0228

$P(X < 55) \approx 2{,}28\%$

alternativ: $P(X < 55) = \dfrac{1}{5\sqrt{2\pi}} \displaystyle\int_{0}^{55} e^{-\left(\frac{x-65}{5}\right)^2} dx \approx 2{,}28\%$

normCdf(0, 55, 65, 5) 0.0228

$P(X < 55) \approx 2{,}28\%$

Antwort: Die Wahrscheinlichkeit, dass der Patient jünger als 55 Jahre ist, beträgt etwa 2,28%.

Hinweis: Im mathematischen Kontext der Normalverteilung muss die untere Grenze $-\infty$ eingesetzt werden. Wer im Sachkontext der Aufgabenstellung arbeitet, kann hier 0 als untere Grenze einsetzen.

zu (4): In einer σ-Umgebung von $1{,}6\sigma$ um den Erwartungswert μ ergibt sich die Wahrscheinlichkeit P zu etwa 89%.

$P(\mu - 1{,}6\sigma \leqslant X \leqslant \mu + 1{,}6\sigma) \approx 89\%$

$\mu - 1{,}6\sigma \leqslant X \leqslant \mu + 1{,}6\sigma$

$57 \leqslant X \leqslant 73$

Antwort: Das Intervall für die Wahrscheinlichkeit, dass das Alter der Patientin in einem Intervall bei 89% liegt, erhält man einen 57-Jährigen als jüngsten und einen 73-Jährigen als ältesten Patienten.

Hinweis: Die verwendeten Befehle müssen vorab im Unterricht eingeführt worden sein. Wird normCdf definiert, so kann auf auch beim Operator *berechnen* auf die ausführliche Integralschreibweise verzichtet werden.

$$\text{normCdf}(a,\ b,\ \mu,\ \sigma) := \frac{1}{\sigma\sqrt{2\pi}} \int_a^b \mathrm{e}^{-\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Bestimmung von Konfidenzintervallen

Aufgabenstellung

Das niederländische Ministerium will Bewohnern von Seniorenstädten Steuererleichterungen in Aussicht stellen. Anhand einer Umfrage unter Senioren soll untersucht werden, ob diese Ankündigung auf den Anteil der Senioren Einfluss hat, die sich für die Seniorenstadt interessieren.
Bei einer Befragung der Tageszeitung bekunden 209 von 540 Senioren ihr Interesse an einer Seniorenstadt. Kann man hieraus schließen, dass der Anteil der Interessierten zugenommen hat? Bestimmen bzw. berechnen Sie dazu ein 95%-Konfidenzintervall und vergleichen Sie dieses mit der Angabe, dass 40% der Senioren Interesse bekundet haben.

Erwartungshorizont

X: Anzahl der interessierten Senioren.
X ist binomialverteilt.

gegeben: $n = 540$; $k = 209$
Die Wahrscheinlichkeit P des gesuchten Intervalls beträgt 95%.
(alternativ: $\mathrm{P}(\mu – 1{,}96\sigma \leqslant \mathrm{X} \leqslant \mu + 1{,}96\sigma) = 0{,}95$)

gesucht: Konfidenzintervall $[p_{\min};\ p_{\max}]$

Ansatz: $\mathrm{P}(\mu – 1{,}96\sigma \leqslant \mathrm{X} \leqslant \mu + 1{,}96\sigma) = 0{,}95$
In einer σ-Umgebung von $1{,}96\sigma$ um den Erwartungswert μ ergibt sich die Wahrscheinlichkeit P zu etwa 95%.

$$\left|\mathrm{X} - n \cdot p\right| \leqslant 1{,}96 \cdot \sqrt{n \cdot p \cdot (1-p)}$$

$\text{solve}(\text{abs}(209–540p) \leqslant 1{,}96 \cdot \sqrt{540 \cdot p \cdot (1-p)}\ ,p)$ $\qquad 0.3469 \leqslant p \leqslant 0.4288$

$p_{\min} = 0{,}3469$ und $p_{\max} = 0{,}4288$

Hinweis: Es kann anstelle der Variablen n, μ und σ bereits im Ansatz der entsprechende Wert der Variabeln eingesetzt werden.

Antwort: Die Wahrscheinlichkeit $p = 0{,}4$ aus dem Aufgabetext ist mit dem berechneten Konfidenzintervall verträglich. Deshalb kann man nicht daraus schließen, dass sich das Interesse der Senioren verändert hat.